



Europäisches Patentamt  
European Patent Office  
Office européen des brevets

(11) Publication number:

0 399 666  
A1

(12)

## EUROPEAN PATENT APPLICATION

(21) Application number: 90304575.5

(51) Int. Cl.<sup>5</sup>: C12N 15/62, C07K 13/00,  
C12P 21/02

(22) Date of filing: 26.04.90

The title of the invention has been amended  
(Guidelines for Examination in the EPO, A-III,  
7.3).

(30) Priority: 29.04.89 GB 8909919

(43) Date of publication of application:  
28.11.90 Bulletin 90/48

(84) Designated Contracting States:  
ES GR

(71) Applicant: Delta Biotechnology Limited  
Castle Court, Castle Boulevard  
Nottingham NG7 1FD(GB)

(72) Inventor: Ballance, David James  
11 South Road  
West Bridgford, Nottingham NG2 7AG(GB)

(74) Representative: Bassett, Richard Simon et al  
ERIC POTTER & CLARKSON St. Mary's Court  
St. Mary's Gate  
Nottingham NG1 1LE(GB)

(54) **Fusion proteins containing N-terminal fragments of human serum albumin.**

(57) A fusion polypeptide comprising, as at least part of the N-terminal portion thereof, an N-terminal portion of HSA or a variant thereof and, as at least part of the C-terminal portion thereof, another polypeptide except that, when the said N-terminal portion of HSA is the 1-n portion where n is 369 to 419 or a variant thereof then the said polypeptide is one of various specified entities, including the 585 to 1578 portion of human fibronectin or a variant thereof.

The HSA-like portion may have additional N-terminal residues, such as secretion leader sequences (signal sequences). The C-terminal portion is preferably the 585-1578 portion of human plasma fibronectin. The N-terminal and C-terminal portions may be cleavable to yield the isolated C-terminal portion, with the N-terminal portion having served to facilitate secretion from the host.

EP 0 399 666 A1

## POLYPEPTIDES

The present invention relates to fusion polypeptides where two individual polypeptides or parts thereof are fused to form a single amino acid chain. Such fusion may arise from the expression of a single continuous coding sequence formed by recombinant DNA techniques.

5 Fusion polypeptides are known, for example those where a polypeptide which is the ultimately desired product of the process is expressed with an N-terminal "leader sequence" which encourages or allows secretion of the polypeptide from the cell. An example is disclosed in EP-A-116 201 (Chiron).

10 Human serum albumin (HSA) is a known protein found in the blood. EP-A-147 198 (Delta Biotechnology) discloses its expression in a transformed host, in this case yeast. Our earlier application EP-A-322 094 discloses N-terminal fragments of HSA, namely those consisting of residues 1-n where n is 369 to 419, which have therapeutic utility. The application also mentions the possibility of fusing the C-terminal residue 15 of such molecules to other, unnamed, polypeptides.

One aspect of the present invention provides a fusion polypeptide comprising, as at least part of the N-terminal portion thereof, an N-terminal portion of HSA or a variant thereof and, as at least part of the C-terminal portion thereof, another polypeptide except that, when the said N-terminal portion of HSA is the 1-n portion where n is 369 to 419 or a variant thereof then the said polypeptide is (a) the 585 to 1578 portion of 20 human fibronectin or a variant thereof, (b) the 1 to 368 portion of CD4 or a variant thereof, (c) platelet derived growth factor, or a variant thereof, (d) transforming growth factor, or a variant thereof, (e) the 1-261 portion of mature human plasma fibronectin or a variant thereof, (f) the 278-578 portion of mature human plasma fibronectin or a variant thereof, (g) the 1-272 portion of mature human von Willebrand's Factor or a variant thereof, or (h) alpha-1-antitrypsin or a variant thereof.

The N-terminal portion of HSA is preferably the said 1-n portion, the 1-177 portion (up to and including the cysteine), the 1-200 portion (up to but excluding the cysteine) or a portion intermediate 1-177 and 1-200.

25 The term "human serum albumin" (HSA) is intended to include (but not necessarily to be restricted to) known or yet-to-be-discovered polymorphic forms of HSA. For example, albumin Naskapi has Lys-372 in place of Glu-372 and pro-albumin Christchurch has an altered pro-sequence. The term "variants" is intended to include (but not necessarily to be restricted to) minor artificial variations in sequence (such as 30 molecules lacking one or a few residues, having conservative substitutions or minor insertions of residues, or having minor variations of amino acid structure). Thus polypeptides which have 80%, preferably 85%, 90%, 95% or 99%, homology with HSA are deemed to be "variants". It is also preferred for such variants to be physiologically equivalent to HSA; that is to say, variants preferably share at least one pharmacological 35 utility with HSA. Furthermore, any putative variant which is to be used pharmacologically should be non-immunogenic in the animal (especially human) being treated.

40 Conservative substitutions are those where one or more amino acids are substituted for others having similar properties such that one skilled in the art of polypeptide chemistry would expect at least the secondary structure, and preferably the tertiary structure, of the polypeptide to be substantially unchanged. For example, typical such substitutions include asparagine for glutamine, serine for asparagine and arginine for lysine. Variants may alternatively, or as well, lack up to ten (preferably only one or two) intermediate 45 amino acid residues (ie not at the termini of the said N-terminal portion of HSA) in comparison with the corresponding portion of natural HSA; preferably any such omissions occur in the 100 to 369 portion of the molecule (relative to mature HSA itself) (if present). Similarly, up to ten, but preferably only one or two, amino acids may be added, again in the 100 to 369 portion for preference (if present). The term "physiologically functional equivalents" also encompasses larger molecules comprising the said sequence plus a further sequence at the N-terminal (for example, pro-HSA, pre-pro-HSA and met-HSA).

45 Clearly, the said "another polypeptide" in the fusion compounds of the invention cannot be the remaining portion of HSA, since otherwise the whole polypeptide would be HSA, which would not then be a "fusion polypeptide".

Even when the HSA-like portion is not the said 1-n portion of HSA, it is preferred for the non-HSA portion to be one of the said (a) to (h) entities.

50 The 1 to 368 portion of CD4 represents the first four disulphide-linked immunoglobulin-like domains of the human T lymphocyte CD4 protein, the gene for and amino acid sequence of which are disclosed in D. Smith *et al* (1987) *Science* 328, 1704-1707. It is used to combat HIV infections.

The sequence of human platelet-derived growth factor (PDGF) is described in Collins *et al* (1985) *Nature* 316, 748-750. Similarly, the sequence of transforming growth factors  $\beta$  (TGF- $\beta$ ) is described in Derynck *et al* (1985) *Nature* 316, 701-705. These growth factors are useful for wound-healing.

A cDNA sequence for the 1-261 portion of Fn was disclosed in EP-A-207 751 (obtained from plasmid pFH6 with endonuclease PvuII). This portion binds fibrin and can be used to direct fused compounds to blood clots.

5 A cDNA sequence for the 278-578 portion of Fn, which contains a collagen-binding domain, was disclosed by R.J. Owens and F.E. Baralle in 1986 E.M.B.O.J. 5, 2825-2830. This portion will bind to platelets.

The 1-272 portion of von Willebrand's Factor binds and stabilises factor VIII. The sequence is given in Bontham et al, Nucl. Acids Res. 14, 7125-7127.

10 Variants of alpha-1-antitrypsin include those disclosed by Rosenburg et al (1984) Nature 312, 77-80. In particular, the present invention includes the Pittsburgh variant (Met<sup>358</sup> is mutated to Arg) and the variant where Pro<sup>357</sup> and Met<sup>358</sup> are mutated to alanine and arginine respectively. These compounds are useful in the treatment of septic shock and lung disorders.

15 Variants of the non-HSA portion of the polypeptides of the invention include variations as discussed above in relation to the HSA portion, including those with conservative amino acid substitutions, and also homologues from other species.

The fusion polypeptides of the invention may have N-terminal amino acids which extend beyond the portion corresponding to the N-terminal portion of HSA. For example, if the HSA-like portion corresponds to an N-terminal portion of mature HSA, then pre-, pro-, or pre-pro sequences may be added thereto, for example the yeast alpha-factor leader sequence. The fused leader portions of WO 90/01063 may be used. 20 The polypeptide which is fused to the HSA portion may be a naturally-occurring polypeptide, a fragment thereof or a novel polypeptide, including a fusion polypeptide. For example, in Example 3 below, a fragment of fibronectin is fused to the HSA portion via a 4 amino acid linker.

It has been found that the amino terminal portion of the HSA molecule is so structured as to favour particularly efficient translocation and export of the fusion compounds of the invention in eukaryotic cells.

25 A second aspect of the invention provides a transformed host having a nucleotide sequence so arranged as to express a fusion polypeptide as described above. By "so arranged", we mean, for example, that the nucleotide sequence is in correct reading frame with an appropriate RNA polymerase binding site and translation start sequence and is under the control of a suitable promoter. The promoter may be homologous with or heterologous to the host. Downstream (3') regulatory sequences may be included if 30 desired, as is known. The host is preferably yeast (for example Saccharomyces spp., e.g. S. cerevisiae; Kluyveromyces spp., e.g. K. lactis; Pichia spp.; or Schizosaccharomyces spp., e.g. S. pombe) but may be any other suitable host such as E. coli, B. subtilis, Aspergillus spp., mammalian cells, plant cells or insect cells.

35 A third aspect of the invention provides a process for preparing a fusion polypeptide according to the first aspect of the invention by cultivation of a transformed host according to the second aspect of the invention, followed by separation of the fusion polypeptide in a useful form.

A fourth aspect of the invention provides therapeutic methods of treatment of the human or other animal body comprising administration of such a fusion polypeptide.

40 In the methods of the invention we are particularly concerned to improve the efficiency of secretion of useful therapeutic human proteins from yeast and have conceived the idea of fusing to amino-terminal portions of HSA those proteins which may ordinarily be only inefficiently secreted. One such protein is a potentially valuable wound-healing polypeptide representing amino acids 585 to 1578 of human fibronectin (referred to herein as Fn 585-1578). As we have described in a separate application (filed simultaneously herewith) this molecule contains cell spreading, chemotactic and chemokinetic activities useful in healing 45 wounds. The fusion polypeptides of the present invention wherein the C-terminal portion is Fn 585-1578 can be used for wound healing applications as biosynthesised, especially where the hybrid human protein will be topically applied. However, the portion representing amino acids 585 to 1578 of human fibronectin can if desired be recovered from the fusion protein by preceding the first amino acid of the fibronectin portion by amino acids comprising a factor X cleavage site. After isolation of the fusion protein from culture 50 supernatant, the desired molecule is released by factor X cleavage and purified by suitable chromatography (e.g. ion-exchange chromatography). Other sites providing for enzymatic or chemical cleavage can be provided, either by appropriate juxtaposition of the N-terminal and C-terminal portions or by the insertion therebetween of an appropriate linker.

55 At least some of the fusion polypeptides of the invention, especially those including the said CD4 and vWF fragments, PDGF and  $\alpha_1$ AT, also have an increased half-life in the blood and therefore have advantages and therapeutic utilities themselves, namely the therapeutic utility of the non-HSA portion of the molecule. In the case of  $\alpha_1$ AT and others, the compound will normally be administered as a one-off dose or only a few doses over a short period, rather than over a long period, and therefore the compounds are less

likely to cause an immune response.

#### EXAMPLES : SUMMARY

5

Standard recombinant DNA procedures were as described by Maniatis et al (1982 and recent 2nd edition) unless otherwise stated. Construction and analysis of phage M13 recombinant clones was as described by Messing (1983) and Sanger et al (1977).

10 DNA sequences encoding portions of human serum albumin used in the construction of the following molecules are derived from the plasmids mHOB12 and pDBD2 (EP-A-322 094, Delta Biotechnology Ltd, relevant portions of which are reproduced below) or by synthesis of oligonucleotides equivalent to parts of this sequence. DNA sequences encoding portions of human fibronectin are derived from the plasmid pFHDEL1, or by synthesis of oligonucleotides equivalent to parts of this sequence. Plasmid pFHDEL1, which contains the complete human cDNA encoding plasma fibronectin, was obtained by ligation of DNA 15 derived from plasmids pFH6, 16, 54, 154 and 1 (EP-A-207 751; Delta Biotechnology Ltd).

This DNA represents an mRNA variant which does not contain the 'ED' sequence and had an 89-amino acid variant of the III-CS region (R.J. Owens, A.R. Kornblith and F.E. Baralle (1986) Oxford Surveys on Eukaryotic Genes 3 141-160). The map of this vector is disclosed in Fig. 11 and the protein sequence of the mature polypeptide produced by expression of this cDNA is shown in Fig. 5.

20 Oligonucleotides were synthesised on an Applied Biosystems 380B oligonucleotide synthesiser according to the manufacturer's recommendations (Applied Biosystems, Warrington, Cheshire, UK).

An expression vector was constructed in which DNA encoding the HSA secretion signal and mature HSA up to and including the 387th amino acid, leucine, fused in frame to DNA encoding a segment of human fibronectin representing amino acids 585 to 1578 inclusive, was placed downstream of the hybrid 25 promoter of EP-A-258 067 (Delta Biotechnology), which is a highly efficient galactose-inducible promoter functional in *Saccharomyces cerevisiae*. The codon for the 1578th amino acid of human fibronectin was directly followed by a stop codon (TAA) and then the *S. cerevisiae* phosphoglycerate kinase (PGK) gene transcription terminator. This vector was then introduced into *S. cerevisiae* by transformation, wherein it directed the expression and secretion from the cells of a hybrid molecule representing the N-terminal 387 30 amino acids of HSA C-terminally fused to amino acids 585 to 1578 of human fibronectin.

In a second example a similar vector is constructed so as to enable secretion by *S. cerevisiae* of a hybrid molecule representing the N-terminal 195 amino acids of HSA C-terminally fused to amino acids 585 to 1578 of human fibronectin.

Aspects of the present invention will now be described by way of example and with reference to the 35 accompanying drawings, in which:

Figure 1 (on two sheets) depicts the amino acid sequence currently thought to be the most representative of natural HSA, with (boxed) the alternative C-termini of HSA(1-n);

Figure 2 (on two sheets) depicts the DNA sequence coding for mature HSA, wherein the sequence included in Linker 3 is underlined;

40 Figure 3 illustrates, diagrammatically, the construction of mHOB16;  
 Figure 4 illustrates, diagrammatically, the construction of pHOB31;  
 Figure 5 (on 6 sheets) illustrates the mature protein sequence encoded by the Fn plasmid pFHDEL1;  
 Figure 6 illustrates Linker 5, showing the eight constituent oligonucleotides;  
 Figure 7 shows schematically the construction of plasmid pDBDF2;  
 45 Figure 8 shows schematically the construction of plasmid pDBDF5;  
 Figure 9 shows schematically the construction of plasmid pDBDF9;  
 Figure 10 shows schematically the construction of plasmid DBDF12, using plasmid pFHDEL1; and  
 Figure 11 shows a map of plasmid pFHDEL1.

50

#### EXAMPLE 1 : HSA 1-387 FUSED TO Fn 585-1578

The following is an account of a preparation of plasmids comprising sequences encoding a portion of HSA, as is disclosed in EP-A-322 094.

55 The human serum albumin coding sequence used in the construction of the following molecules is derived from the plasmid M13mp19.7 (EP-A-201 239, Delta Biotechnology Ltd.) or by synthesis of oligonucleotides equivalent to parts of this sequence. Oligonucleotides were synthesised using phosphoramidite chemistry on an Applied Biosystems 380B oligonucleotide synthesizer according to the

manufacturer's recommendations (AB Inc., Warrington, Cheshire, England).

An oligonucleotide was synthesised (Linker A) which represented a part of the known HSA coding sequence (Figure 2) from the PstI site (1235-1240, Figure 2) to the codon for valine 381 wherein that codon was changed from GTG to GTC:

5

Linker 1

10

		D	P	H	E	C	Y
15	5'	GAT	CCT	CAT	GAA	TGC	TAT
	3' ACGT	CTA	GGA	GTA	CTT	ACG	ATA

1247

20

	A	K	V	F	D	E	F	K
25	GCC	AAA	GTG	TTC	GAT	GAA	TTT	AAA
	CGG	TTT	CAC	AAG	CTA	CTT	AAA	TTT

1267

30

P	L	V'
CTT	GTC	3'
GGA	CAG	5'

35

Linker 1 was ligated into the vector M13mp19 (Norlander et al, 1983) which had been digested with PstI and HincII and the ligation mixture was used to transfet E.coli strain XL1-Blue (Stratagene Cloning Systems, San Diego, CA). Recombinant clones were identified by their failure to evolve a blue colour on medium containing the chromogenic indicator X-gal (5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactoside) in the present of IPTG (isopropylthio- $\beta$ -galactoside). DNA sequence analysis of template DNA prepared from bacteriophage particles of recombinant clones identified a molecule with the required DNA sequence, designated mHOB12 (Figure 3). M13mp19.7 consists of the coding region of mature HSA in M13mp19 (Norlander et al, 1983) such that the codon for the first amino acid of HSA, GAT, overlaps a unique XbaI site thus:

50

	Asp	Ala	
5'	C T C G A G A T G C A		3'
3'	G A G C T C T A C G T		5'

55

XbaI

(EP-A-210 239). M13mp19.7 was digested with Xhol and made flush-ended by S1-nuclease treatment and was then ligated with the following oligonucleotide (Linker 2):

5 Linker 2

10 5' T C T T T A T C C A A G C T T G G A T A A A A G A 3'  
 3' A G A A A A T A G G T T C G A A C C T A T T T C T 5'  
HindIII

15

The ligation mix was then used to transfect E.coli XL1-Blue and template DNA was prepared from several plaques and then analysed by DNA sequencing to identify a clone, pDBD1 (Figure 4), with the correct sequence.

20 A 1.1 kb HindIII to PstI fragment representing the 5' end of the HSA coding region and one half of the inserted oligonucleotide linker was isolated from pDBD1 by agarose gel electrophoresis. This fragment was then ligated with double stranded mHOB12 previously digested with HindIII and PstI and the ligation mix was then used to transfect E.coli XL1-Blue. Single stranded template DNA was prepared from mature bacteriophage particles of several plaques. The DNA was made double stranded *in vitro* by extension from annealed sequencing primer with the Klenow fragment of DNA polymerase I in the presence of deoxyribonucleoside triphosphates. Restriction enzyme analysis of this DNA permitted the identification of a clone with the correct configuration, mHOB15 (Figure 4).

25 The following oligonucleotide (Linker 3) represents from the codon for the 382nd amino acid of mature HSA (glutamate, GAA) to the codon for lysine 389 which is followed by a stop codon (TAA) and a HindIII site and then a BamHI cohesive end:

30 Linker 3

35

E E P Q N L I K J

3' GAA GAG CCT CAG AAT TTA ATC AAA TAA GCTTG 3'  
 40 3' CTT CTC GGA GTC TTA AAT TAG TTT ATT CGAACCTAG 5'

45 This was ligated into double stranded mHOB15, previously digested with HincII and BamHI. After ligation, the DNA was digested with HincII to destroy all non-recombinant molecules and then used to transfect E.coli XL1-Blue. Single stranded DNA was prepared from bacteriophage particles of a number of clones and subjected to DNA sequence analysis. One clone having the correct DNA sequence was designated mHOB16 (Figure 4).

50 A molecule in which the mature HSA coding region was fused to the HSA secretion signal was created by insertion of Linker 4 into BamHI and Xhol digested M13mp19.7 to form pDBD2 (Figure 4).

Linker 4

55

		M	K	W	V	S	F		
5	5'	GATCC	ATG	AAG	TGG	GTA	AGC	TTT	
		G	TAC	TTC	ACC	CAT	TCG	AAA	
10		I	S	L	L	F	L	F	S
		ATT	TCC	CTT	CTT	TTT	CTC	TTT	AGC
15		TAA	AGG	GAA	GAA	AAA	GAG	AAA	TCG
20		S	A	Y	S	R	G	V	F
		TCG	GCT	TAT	TCC	AGG	GGT	GTG	TTT
25		AGC	CGA	ATA	AGG	TCC	CCA	CAC	AAA
		R	R						
30		CG	3'						
		GCAGCT	5'						

35 In this linker the codon for the fourth amino acid after the initial methionine, ACC for threonine in the HSA pre-pro leader sequence (Lawn *et al.*, 1981), has been changed to AGC for serine to create a HindIII site.

40 A sequence of synthetic DNA representing a part of the known HSA coding sequence (Lawn *et al.*, 1981) (amino acids 382 to 387, Fig. 2), fused to part of the known fibronectin coding sequence (Kornblith *et al.*, 1985) (amino acids 585 to 640, Fig. 2), was prepared by synthesising six oligonucleotides (Linker 5, Fig. 6). The oligonucleotides 2, 3, 4, 6, 7 and 8 were phosphorylated using T4 polynucleotide kinase and then the oligonucleotides were annealed under standard conditions in pairs, i.e. 1+8, 2+7, 3+6 and 4+5. The annealed oligonucleotides were then mixed together and ligated with mHOB12 which had previously been digested with the restriction enzymes HincII and EcoRI. The ligation mixture was then used to transfet 45 E.coli XL1-Blue (Stratagene Cloning Systems, San Diego, CA). Single stranded template DNA was then prepared from mature bacteriophage particles derived from several independent plaques and then was analysed by DNA sequencing. A clone in which a linker of the expected sequence had been correctly inserted into the vector was designated pDBDF1 (Fig. 7). This plasmid was then digested with PstI and EcoRI and the approx. 0.24kb fragment was purified and then ligated with the 1.29kb BamHI-PstI fragment 50 of pDBD2 (Fig. 7) and BamHI + EcoRI digested pUC19 (Yanisch-Perron, *et al.*, 1985) to form pDBDF2 (Fig. 7).

55 A plasmid containing a DNA sequence encoding full length human fibronectin, pFHDEL1, was digested with EcoRI and XbaI and a 0.77kb EcoRI-XbaI fragment (Fig. 8) was isolated and then ligated with EcoRI and SacI digested M13 mp18 (Norrander *et al.*, 1983) to form pDBDF3 (Fig. 8).

The following oligonucleotide linker (Linker 6) was synthesised, representing from the PstI site at 4784-4791 of the fibronectin sequence of EP-A-207 751 to the codon for tyrosine 1578 (Fig. 5) which is followed by a stop codon (TAA), a HindIII site and then a BamHI cohesive end:

## Linker 6

5 G P D Q T E M T I E G L  
GGT CCA GAT CAA ACA GAA ATG ACT ATT GAA GGC TTG  
10 A CGT CCA GGT CTA GTT TGT CTT TAC TGA TAA CTT CCG AAC

	Q	P	T	V	E	Y	Stop	
15	CAG	CCC	ACA	GTG	GAG	TAT	TAA	GCTTG
	GTC	GGG	TGT	CAC	CTC	ATA	ATT	CGAACCTAG

20 This linker was then ligated with PstI and HindIII digested pDBDF3 to form pDBDF4 (Fig. 8). The following DNA fragments were then ligated together with BglII digested pKV50 (EP-A-258 067) as shown in Fig. 8: 0.68kb EcoRI-BamHI fragment of pDBDF4, 1.5kb BamHI-StuI fragment of pDBDF2 and the 2.2kb StuI-EcoRI fragment of pFHDEL1. The resultant plasmid pDBDF5 (Fig. 8) includes the promoter of EP-A-258 067 to direct the expression of the HSA secretion signal fused to DNA encoding amino acids 1-387 of mature HSA, in turn fused directly and in frame with DNA encoding amino acids 585-1578 of human fibronectin, after which translation would terminate at the stop codon TAA. This is then followed by the *S.cerevisiae* PGK gene transcription terminator. The plasmid also contains sequences which permit selection and maintenance in *Escherichia coli* and *S.cerevisiae* (EP-A-258 067).

25 30 This plasmid was introduced into *S.cerevisiae* S150-2B (leu2-3 leu2-112 ura3-52 trp1-289 his3-1) by standard procedures (Beggs, 1978). Transformants were subsequently analysed and found to produce the HSA-fibronectin fusion protein.

35 EXAMPLE 2 : HSA 1-195 FUSED TO Fn 585-1578

In this second example the first domain of human serum albumin (amino acids 1-195) is fused to amino acids 585-1578 of human fibronectin.

The plasmid pDBD2 was digested with BamH I and BglII and the 0.79kb fragment was purified and then ligated with BamH I-digested M13mp19 to form pDBDF6 (Fig. 6). The following oligonucleotide:  
 5'-C C A A A G C T C G A G G A A C T T C G-3'  
 was used as a mutagenic primer to create a XbaI site in pDBDF6 by in vitro mutagenesis using a kit supplied by Amersham International PLC. This site was created by changing base number 696 of HSA from a T to a G (Fig. 2). The plasmid thus formed was designated pDBDF7 (Fig. 9). The following linker was then synthesised to represent from this newly created XbaI site to the codon for lysine 195 of HSA (AAA) and then from the codon for isoleucine 585 of fibronectin to the ends of oligonucleotides 1 and 8 shown in Fig. 6.

50 Linker 7

D E L ' R D E G K A S S A K

5 TC GAT GAA CTT CGG GAT GAA GGG AAG GCT TCG TCT GCC AAA

A CTT GAA GCC CTA CTT CCC TTC CGA AGC AGA CGG TTT

10 I T E T P S Q P N S H

ATC ACT GAG ACT CCG AGT CAG C

15 TAG TGA CTC TGA GGC TCA GTC GGG TTG AGG GTG G

This linker was ligated with the annealed oligonucleotides shown in Fig. 3, i.e. 2+7, 3+6 and 4+5 together with Xhol and EcoRI digested pDBDF7 to form pDBDF8 (Fig. 9). Note that in order to recreate the original HSA DNA sequence, and hence amino acid sequence, insertion of linker 7 and the other oligonucleotides into pDBDF7 does not recreate the Xhol site.

The 0.83kb BamHi-Stu fragment of pDBDF8 was purified and then was ligated with the 0.68kb EcoRI-BamHI fragment of pDBDF2 and the 2.22kb Stu-EcoRI fragment of pFHDEL1 into BglII-digested pKV50 to form pDBDF9 (Fig. 9). This plasmid is similar to pDBDF5 except that it specifies only residues 1-195 of HSA rather than 1-387 as in pDBDF5.

When introduced into S.cerevisiae S150-2B as above, the plasmid directed the expression and secretion of a hybrid molecule composed of residues 1-195 of HSA fused to residues 585-1578 of fibronectin.

30 EXAMPLE 3 : HSA 1-387 FUSED TO Fn 585-1578, AS CLEAVABLE MOLECULE

In order to facilitate production of large amounts of residues 585-1578 of fibronectin, a construct was made in which DNA encoding residues 1-387 of HSA was separated from DNA encoding residues 585-1578 of fibronectin by the sequence

35

I E G R

40 ATT GAA GGT AGA

TAA CTT CCA TCT

45 which specifies the cleavage recognition site for the blood clotting Factor X. Consequently the purified secreted product can be treated with Factor X and then the fibronectin part of the molecule can be separated from the HSA part.

To do this two oligonucleotides were synthesised and then annealed to form Linker 8.

50

Linker 8

55

	E	E	P	Q	N	L	I	E	G
5	GAA	GAG	CCT	CAG	AAT	TTA	ATT	GAA	GGT
	CTT	CTC	GGA	GTC	TTA	AAT	CAA	CTT	CCA
10	R	I	T	E	T	P	S	Q	P
	AGA	ATC	ACT	GAG	ACT	CCG	AGT	CAG	C
15	TCT	TAG	TGA	CTC	TGA	GGC	TCA	GTC	GGG
	N	S	H						
20	TTG	AGG	GTG	G					

25 This linker was then ligated with the annealed oligonucleotides shown in Fig. 6, i.e. 2+7, 3+6 and 4+5 into HincII and EcoRI digested mHOB12, to form pDBDF10 (Fig. 7). The plasmid was then digested with PstI and EcoRI and the roughly 0.24kb fragment was purified and then ligated with the 1.29kb BamHi-PstI fragment of pDBD2 and BamHi and EcoRI digested pUC19 to form pDBDF11 (Fig. 10).

30 The 1.5kb BamHi-StuI fragment of pDBDF11 was then ligated with the 0.68kb EcoRI-BamH1 fragment of pDBDF4 and the 2.22kb StuI-EcoRI fragment of pFHDEL1 into BglII-digested pKV50 to form pDBDF12 (Fig. 10). This plasmid was then introduced into S.cerevisiae S150-2B. The purified secreted fusion protein was treated with Factor X to liberate the fibronectin fragment representing residues 585-1578 of the native molecule.

35

#### REFERENCES

40 Beggs, J.D. (1978) Nature 275, 104-109  
 Kornblith et al. (1985) EMBO J. 4, 1755-1759  
 Lawn, R.M. et al. (1981) Nucl. Acid. Res. 9, 6103-6114  
 Maniatis, T. et al. (1982) Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.  
 Messing, J. (1983) Methods Enzymol. 101, 20-78  
 45 Norrander, J. et al. (1983) Gene 26, 101-106  
 Sanger, F. et al. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467  
 Yanisch-Perron, C. (1985) Gene 33, 103-119

50 **Claims**

1. A fusion polypeptide comprising, as at least part of the N-terminal portion thereof, an N-terminal portion of HSA or a variant thereof and, as at least part of the C-terminal portion thereof, another polypeptide except that, when the said N-terminal portion of HSA is the 1-n portion where n is 369 to 419 or 55 a variant thereof then the said polypeptide is (a) the 585 to 1578 portion of human fibronectin or a variant thereof, (b) the 1 to 368 portion of CD4 or a variant thereof, (c) platelet derived growth factor or a variant thereof, (d) transforming growth factor  $\beta$  or a variant thereof, (e) the 1-261 portion of mature human plasma fibronectin or a variant thereof, (f) the 278-578 portion of mature human plasma fibronectin or a variant

thereof, (g) the 1-272 portion of mature human von Willebrand's Factor or a variant thereof, or (h) alpha-1-antitrypsin or a variant thereof.

2. A fusion polypeptide according to Claim 1 additionally comprising at least one N-terminal amino acid extending beyond the portion corresponding to the N-terminal portion of HSA.

5 3. A fusion polypeptide according to Claim 1 or 2 wherein there is a cleavable region at the junction of the said N-terminal or C-terminal portions.

4. A fusion polypeptide according to any one of the preceding claims wherein the said C-terminal portion is the 585 to 1578 portion of human plasma fibronectin or a variant thereof.

10 5. A transformed or transfected host having a nucleotide sequence so arranged as to express a fusion polypeptide according to any one of the preceding claims.

6. A process for preparing a fusion polypeptide by cultivation of a host according to Claim 5, followed by separation of the fusion polypeptide in a useful form.

7. A fusion polypeptide according to any one of Claims 1 to 4 for use in therapy.

15

20

25

30

35

40

45

50

55

FIGURE 1

10 Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys  
30 Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln Cys Pro Phe Glu Asp His Val  
50 Lys Leu Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu  
70 Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu  
90 Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu  
110 Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val  
130 Asp Val Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr  
150 Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg  
170 Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro  
190 Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys  
210 Ala Ser Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser  
230 Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys  
250 Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu  
270 Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu  
290 Lys Pro Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala  
310 Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala  
330 Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp  
350 Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys  
370 Cys Ala Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu

FIGURE 1 Cont.

390	400
Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu	
410	420
Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln Val Ser	Thr
430	440
Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His	
450	460
Pro Glu Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu	
470	480
Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser	
490	500
Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr Val Pro Lys	
510	520
Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu	
530	540
Arg Gln Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr	
550	560
Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys	
570	580
Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln	
Ala Ala Leu Gly Leu	

**FIGURE 2** DNA sequence coding for mature HSA

10            20            30            40            50            60            70            80  
 GATGCACACACAAGAGTGAGGTTGCTCATCGGTTAAAGATTGGAGAAGAAAATTCAAAGCCTGGTGTGATTGCCTT  
 D A H K S E V A H R F K D L G E E N F K A L V L I A F  
  
 90            100          110          120          130          140          150          160  
 TGCTCAGTATCTTCAGCAGTGTCCATTGAAAGATCATGTAAGGAAATTAGTGAATGAAGTAACGTGAATTTGC  
 A Q Y L Q Q C P F E D H V K L V N E V T E F A K T C  
  
 170          180          190          200          210          220          230          240  
 TTGCTGATGAGTCAGCTGAAAATTGTGACAAATCACTTCATACCCCTTTGGAGACAAATTATGCACAGTTGCAACTCCTT  
 V A D E S A E N C D K S L H T L F G D K L C T V A T L  
  
 250          260          270          280          290          300          310          320  
 CGTGAAACCTATGGTGAATGGCTGACTGCTGTGCAAAACAAGAACCTGAGAGAAATGAATGCTTCTTGC  
 R E T Y G E M A D C C A K Q E P E R N E C F L Q H K D  
  
 330          340          350          360          370          380          390          400  
 TGACAACCCAAACCTCCCCGATTGGTGAGACCAGAGGTTGATGTGCACTGCTTTCATGACAATGAAGAGACAT  
 D N P N L P R L V R P E V D V M C T A F H D N E E T  
  
 410          420          430          440          450          460          470          480  
 TTTTGAAAAAAACTTATGAAATTGCCAGAAGACATCCTTACTTTATGCCCGGAACTCCTTTCTTGCTAAAAGG  
 F L K K Y L Y E I A R R H P Y F Y A P E L L F F A K R  
  
 490          500          510          520          530          540          550          560  
 TATAAAAGCTGCTTTACAGAATGTTGCCAGCTGCTGATAAAAGCTGCCCTGCTGTTGCCAAAGCTCGATGA  
 Y K A A F T E C C Q A A D K A A C L L P K L D E L R D  
  
 570          580          590          600          610          620          630          640  
 TGAAGGGAAAGGCTTCGCTCTGCCAACACAGAGACTCAAATGTGCCAGTCTCCAAAAATTGGAGAAAGAGCTTCAA  
 E G K A S S A K Q R L K C A S L Q K F G E R A F K A  
  
 650          660          670          680          690          700          710          720  
 GGGCAGTGGCTCGCTGAGCCAGAGATTCCCAAAGCTGAGTTGCAGAAGTTCCAAGTTAGTGCACAGATCTTAC  
 W A V A R L S Q R F P K A E F A E V S K L V T D L T K  
  
 730          740          750          760          770          780          790          800  
 GTCCACACGGAATGCTGCCATGGAGATCTGCTGAATGTGCTGATGACAGGGCGGACCTTGC  
 V H T E C C H G D L L E C A D D R A D L A K Y I C E N  
  
 810          820          830          840          850          860          870          880  
 TCAGGATTCCGATCTCCAGTAAACTGAAGGAATGCTGTGAAAAACCTCTGTTGGAAAAATCCC  
 Q D S I S S K L K E C C E K P L L E K S H C I A E V  
  
 890          900          910          920          930          940          950          960  
 AAAATGATGAGATGCCGCTGACTTGCCTTCATTAGCTGCTGATTTGTGAAAGTAAGGATGTTGC  
 E N D E M P A D L P S L A A D F V E S K D V C K N Y A  
  
 970          980          990          1000        1010        1020        1030        1040  
 GAGGCAAAGGATGTCTCCTGGCATGTTTGATGAATATGCAAGAAGGC  
 E A K D V F L G M F L Y E Y A R R H P D Y S V V L L L

FIGURE 2 Cont.

1050 1060 1070 1080 1090 1100 1110 1120  
 GAGACTGCCAAGACATATGAAACCACTCTAGAGAAGTGTGCGCTGCAGATCCTCATGAATGCTATGCCAAAGTGT  
 R L A K T Y E T T L E K C C A A A D P H E C Y A K V  
  
 1130 1140 1150 1160 1170 1180 1190 1200  
 TCGATGAATTAAACCTCTGTGGAAGAGCCTCAGAATTAAATCAAACAAAAGTGTGAGCTTTGAGCAGCTGGAGAG  
 F D E F K P L V E E P Q N L I K Q N C E L F E Q L G E  
  
 1210 1220 1230 1240 1250 1260 1270 1280  
 TACAAATTCCAGAATGCGCTATTAGTTCGTTACACCAAGAAAGTACCCCAAGTGTCAACTCCAACCTCTGTAGAGGTCTC  
 Y K F Q N A L L V R Y T K K V P Q V S T P T L V E V S  
  
 1290 1300 1310 1320 1330 1340 1350 1360  
 AAGAAACCTAGGAAAAGTGGGCAGCAAATGTTGAAACATCCTGAAGCAAAAGAATGCCCTGTGCAGAAGACTATCTAT  
 R N L G K V G S K C C K H P E A K R M P C A E D Y L  
  
 1370 1380 1390 1400 1410 1420 1430 1440  
 CCGTGGTCTGAACCAGTTATGTTGATGAGAAAAGCCAGTAAGTGACAGAGTCACAAAATGCTGCACAGAGTCC  
 S V V L N Q L C V L H E K T P V S D R V T K C C T E S  
  
 1450 1460 1470 1480 1490 1500 1510 1520  
 TTGGTGAACAGGCAGCATGCTTCAGCTCTGGAAAGTCGATGAAACATACGTTCCAAAGAGTTAATGCTGAAACATT  
 L V N R R P C F S A L E V D E T Y V P K E F N A E T F  
  
 1530 1540 1550 1560 1570 1580 1590 1600  
 CACCTTCATGCAGATATATGCACACTTCTGAGAAGGAGAGACAAATCAAGAAACAAACTGCACTTGTTGAGCTGTGA  
 T F H A D I C T L S E K E R Q I K K Q T A L V E L V  
  
 1610 1620 1630 1640 1650 1660 1670 1680  
 AACACAAGCCAAGGCAACAAAAGAGCAACTGAAAGCTGTTATGGATGATTCCGAGCTTGAGAGAAGTGCTGCAAG  
 K H K P K A T K E Q L K A V M D D F A A F V E K C C K  
  
 1690 1700 1710 1720 1730 1740 1750 1760  
 GCTGACGATAAGGAGACCTGCTTGCCGAGGAGGGTAAAAAAACTTGTGCTGCAAGTCAGCTGCCCTAGGCTTATAACA  
 A D D K E T C F A E E G K K L V A A S Q A A L G L  
  
 1770 1780  
 TCTACATTAAAAGCATCTCAG

FIGURE 3 Construction of mHOB16

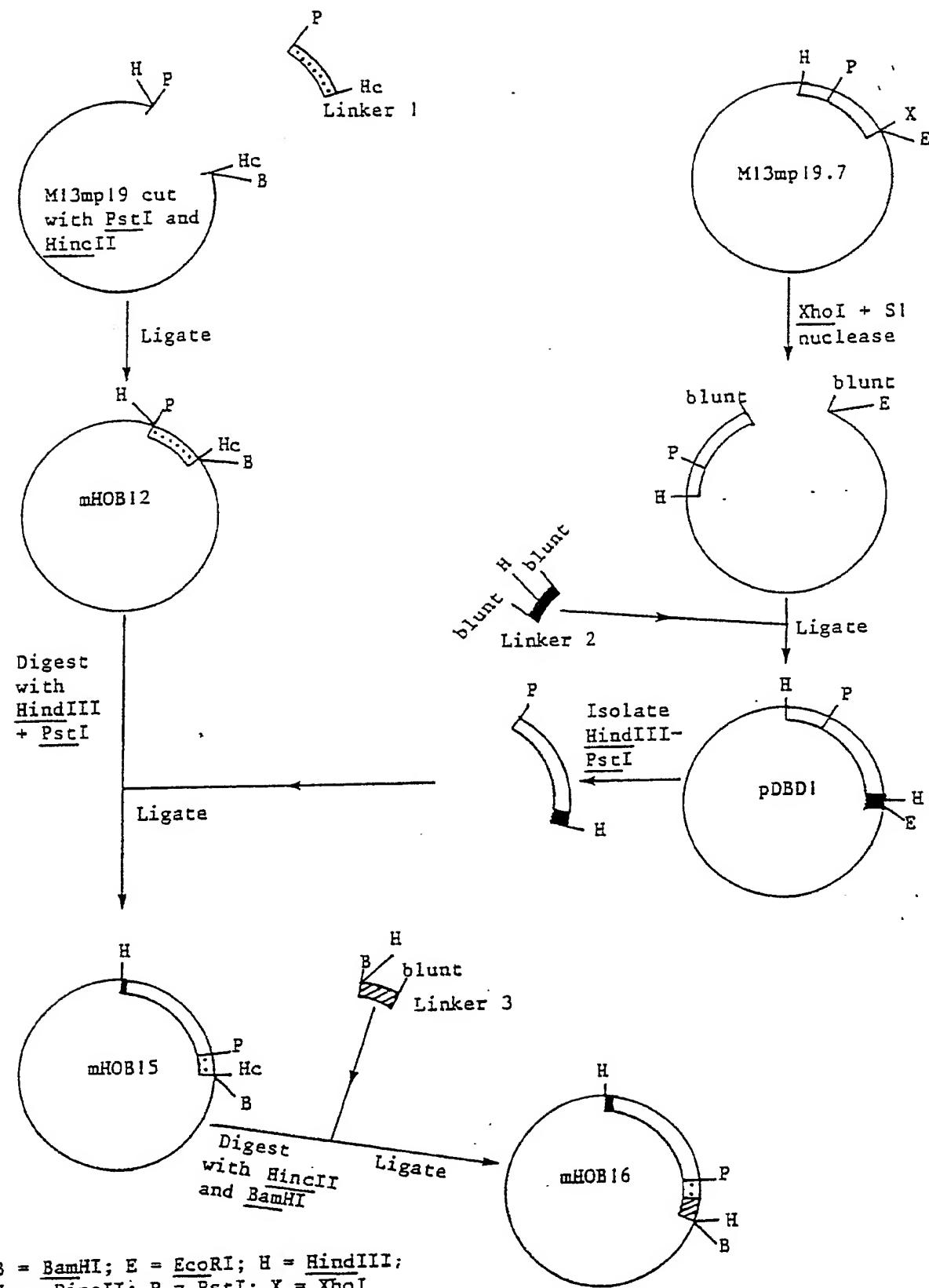
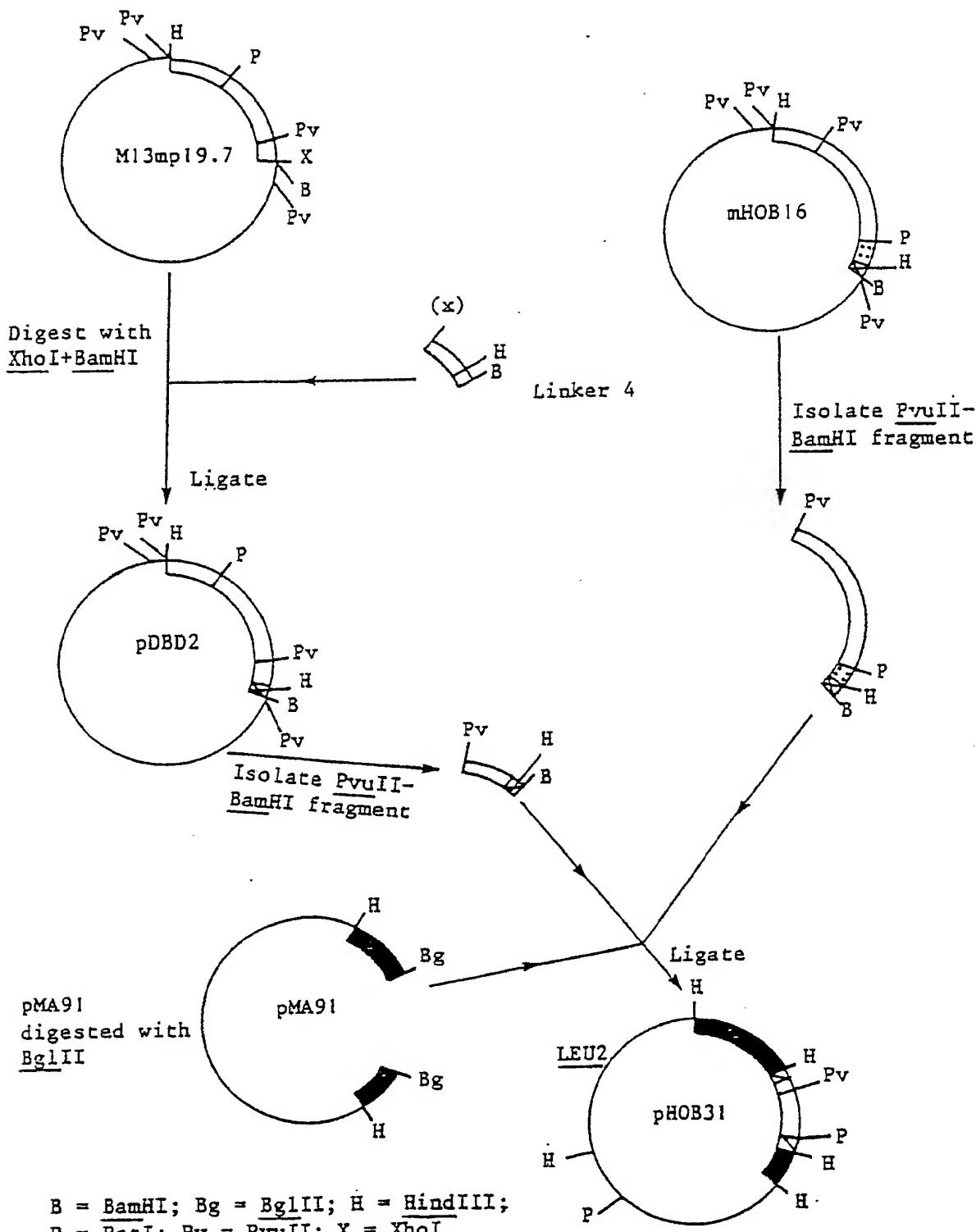


FIGURE 4 Construction of pHOB31



Gln Ala Gln Gln Met Val Gln Pro Gln Ser Pro Val Ala Val Ser Gln Ser Lys Pro Gly 20  
 Cys Tyr Asp Asn Gly Lys His Tyr Gln Ile Asn Gln Gln Trp Glu Arg Thr Tyr Leu Gly 40  
 Asn Val Leu Val Cys Thr Cys Tyr Gly Gly Ser Arg Gly Phe Asn Cys Glu Ser Lys Pro 60  
 Glu Ala Glu Glu Thr Cys Phe Asp Lys Tyr Thr Gly Asn Thr Tyr Arg Val Gly Asp Thr 80  
 Tyr Glu Arg Pro Lys Asp Ser Met Ile Trp Asp Cys Thr Cys Ile Gly Ala Gly Arg Gly 100  
 Arg Ile Ser Cys Thr Ile Ala Asn Arg Cys His Glu Gly Gly Gln Ser Tyr Lys Ile Gly 120  
 Asp Thr Trp Arg Arg Pro His Glu Thr Gly Gly Tyr Met Leu Glu Cys Val Cys Leu Gly 140  
 Asn Gly Lys Gly Glu Trp Thr Cys Lys Pro Ile Ala Glu Lys Cys Phe Asp His Ala Ala 160  
 Gly Thr Ser Tyr Val Val Gly Glu Thr Trp Glu Lys Pro Tyr Gln Gly Trp Met Met Val 180  
 Asp Cys Thr Cys Leu Gly Glu Gly Ser Gly Arg Ile Thr Cys Thr Ser Arg Asn Arg Cys 200  
 Asn Asp Gln Asn Thr Arg Thr Ser Tyr Arg Ile Gly Asp Thr Trp Ser Lys Lys Asn Asn 220  
 Arg Gly Asn Leu Leu Gln Cys Ile Cys Thr Gly Asn Gly Arg Gly Glu Trp Lys Cys Glu 240  
 Arg His Thr Ser Val Gln Thr Thr Ser Ser Gly Ser Gly Pro Phe Thr Asp Val Arg Ala 260  
 Ala Val Tyr Gln Pro Gln Pro His Pro Gln Pro Pro Pro Tyr Gly His Cys Val Thr Asp 280  
 Ser Gly Val Val Tyr Ser Val Gly Met Gln Trp Leu Lys Thr Gln Gly Asn Lys Gln Met 300  
 Leu Cys Thr Cys Leu Gly Asn Gly Val Ser Cys Gln Glu Thr Ala Val Thr Gln Thr Tyr 320  
 Gly Gly Asn Ser Asn Gly Glu Pro Cys Val Leu Pro Phe Thr Tyr Asn Gly Arg Thr Phe 340  
 Tyr Ser Cys Thr Thr Glu Gly Arg Gln Asp Gly His Leu Trp Cys Ser Thr Thr Ser Asn 360  
 Tyr Glu Gln Asp Gln Lys Tyr Ser Phe Cys Thr Asp His Thr Val Leu Val Gln Thr Gin 380  
 Gly Gly Asn Ser Asn Gly Ala Leu Cys His Phe Pro Phe Leu Tyr Asn Asn His Asn Tyr 400  
 Thr Asp Cys Thr Ser Glu Gly Arg Arg Asp Asn Met Lys Trp Cys Gly Thr Thr Gln Asn 420

FIGURE 5A

FNDL1

Tyr Asp Arg Asp Glu Lys Phe Gly Phe Cys Pro Met Ala Ala His Glu Glu Ile Cys Thr 440  
 Tyr Asn Glu Glu Val Met Tyr Arg Ile Gly Asp Glu Trp Asp Lys Glu His Asp Met Gly 450  
 His Met Met Arg Cys Thr Cys Val Gly Asn Gly Arg Gly Glu Trp Thr Cys Tyr Ala Tyr 460  
 Ser Glu Leu Arg Asp Glu Cys Ile Val Asp Asp Ile Thr Tyr Asn Val Asn Asp Thr Phe 470  
 His Lys Arg His Glu Glu Glu His Met Leu Asn Cys Thr Cys Phe Gly Glu Gly Arg Gly 480  
 Arg Trp Lys Cys Asp Pro Val Asp Glu Cys Glu Asp Ser Glu Thr Gly Thr Phe Tyr Glu 490  
 Ile Gly Asp Ser Trp Glu Lys Tyr Val His Gly Val Arg Tyr Glu Cys Tyr Cys Tyr Gly 500  
 Arg Glu Ile Glu Glu Ile His Cys Glu Pro Leu Glu Thr Tyr Pro Ser Ser Gly Pro 510  
 Val Glu Val Phe Ile Thr Glu Thr Pro Ser Glu Pro Asn Ser His Pro Ile Glu Trp Asn 520  
 Ala Pro Glu Pro Ser His Ile Ser Lys Tyr Ile Leu Arg Trp Arg Pro Lys Asn Ser Val 530  
 Glu Arg Trp Lys Glu Ala Thr Ile Pro Gly His Leu Asn Ser Tyr Thr Ile Lys Gly Leu 540  
 Lys Pro Glu Val Val Tyr Glu Glu Glu Ile Ser Ile Glu Glu Tyr Gly His Glu Glu 550  
 Val Thr Arg Phe Asp Phe Thr Thr Ser Thr Ser Thr Pro Val Thr Ser Asn Thr Val 560  
 Thr Gly Glu Thr Thr Pro Phe Ser Pro Leu Val Ala Thr Ser Glu Ser Val Thr Glu Ile 570  
 Thr Ala Ser Ser Phe Val Val Ser Trp Val Ser Ala Ser Asp Thr Val Ser Gly Phe Arg 580  
 Val Glu Tyr Glu Leu Ser Glu Glu Gly Asp Glu Pro Glu Tyr Leu Asp Leu Pro Ser Thr 590  
 Ala Thr Ser Val Asn Ile Pro Asp Leu Leu Pro Glu Arg Lys Tyr Ile Val Asn Val Tyr 600  
 Glu Ile Ser Glu Asp Gly Glu Glu Glu Ser Leu Ile Leu Ser Thr Ser Glu Thr Ala Pro 610  
 Asp Ala Pro Pro Asp Pro Thr Val Asp Glu Val Asp Asp Thr Ser Ile Val Val Arg Thr 620  
 Ser Arg Pro Glu Ala Pro Ile Thr Gly Tyr Arg Ile Val Tyr Ser Pro Ser Val Glu Glu 630  
 Ser Ser Thr Glu Leu Asn Leu Pro Glu Ile Ala Asn Ser Val Thr Leu Ser Asp Leu Glu 640

FNDL1

FIGURE 5B

Pro Gly Val Gln Tyr Asn Ile Thr Ile Tyr Ala Val Glu Glu Asn Gln Glu Ser Thr Pro 860  
 Val Val Ile Gln Gln Glu Thr Thr Gly Thr Pro Arg Ser Asp Thr Val Pro Ser Pro Arg 880  
 Asp Leu Gln Phe Val Glu Val Thr Asp Val Lys Val Thr Ile Met Trp Thr Pro Pro Glu 900  
 Ser Ala Val Thr Gly Tyr Arg Val Asp Val Ile Pro Val Asn Leu Pro Gly Glu His Gly 920  
 Gln Arg Leu Pro Ile Ser Arg Asn Thr Phe Ala Glu Val Thr Gly Leu Ser Pro Gly Val 940  
 Thr Tyr Tyr Phe Lys Val Phe Ala Val Ser His Gly Arg Glu Ser Lys Pro Leu Thr Ala 960  
 Gln Gln Thr Thr Lys Leu Asp Ala Pro Thr Asn Leu Gln Phe Val Asn Glu Thr Asp Ser 980  
 Thr Val Leu Val Arg Trp Thr Pro Pro Arg Ala Gln Ile Thr Gly Tyr Arg Leu Thr Val 1000  
 Gly Leu Thr Arg Arg Gly Gln Pro Arg Gln Tyr Asn Val Gly Pro Ser Val Ser Lys Tyr 1020  
 Pro Leu Arg Asn Leu Gln Pro Ala Ser Glu Tyr Thr Val Ser Leu Val Ala Ile Lys Gly 1040  
 Asn Gln Glu Ser Pro Lys Ala Thr Gly Val Phe Thr Thr Leu Gln Pro Gly Ser Ser Ile 1060  
 Pro Pro Tyr Asn Thr Glu Val Thr Glu Thr Ile Val Ile Thr Trp Thr Pro Ala Pro 1080  
 Arg Ile Gly Phe Lys Leu Gly Val Arg Pro Ser Gln Gly Glu Ala Pro Arg Glu Val 1100  
 Thr Ser Asp Ser Gly Ser Ile Val Val Ser Gly Leu Thr Pro Gly Val Glu Tyr Val Tyr 1120  
 Thr Ile Gln Val Leu Arg Asp Gly Glu Arg Asp Ala Pro Ile Val Asn Lys Val Val 1140  
 Thr Pro Leu Ser Pro Pro Thr Asn Leu His Leu Glu Ala Asn Pro Asp Thr Gly Val Leu 1160  
 Thr Val Ser Trp Glu Arg Ser Thr Thr Pro Asp Ile Thr Gly Tyr Arg Ile Thr Thr Thr 1180  
 Pro Thr Asn Gly Gln Gln Gly Asn Ser Leu Glu Glu Val Val His Ala Asp Gln Ser Ser 1200  
 Cys Thr Phe Asp Asn Leu Ser Pro Gly Leu Glu Tyr Asn Val Ser Val Tyr Thr Val Lys 1220  
 Asp Asp Lys Glu Ser Val Pro Ile Ser Asp Thr Ile Ile Pro Ala Val Pro Pro Pro Thr 1240  
 Asp Leu Arg Phe Thr Asn Ile Gly Pro Asp Thr Met Arg Val Thr Trp Ala Pro Pro Pro 1260

FIGURE 5C

FNOEL1

Ser Ile Asp Leu Thr Asn Phe Leu Val Arg Tyr Ser Pro Val Lys Asn Glu Glu Asn Val 1280  
 Ala Glu Ile Ser Ile Ser Pro Ser Asp Asn Ala Val Val Leu Thr Asn Leu Leu Pro Gly 1300  
 Thr Glu Tyr Val Val Ser Val Ser Ser Val Tyr Glu Glu His Glu Ser Thr Pro Leu Arg 1320  
 Gly Arg Glu Lys Thr Gly Leu Asp Ser Pro Thr Gly Ile Asp Phe Ser Asp Ile Thr Ala 1340  
 Asn Ser Phe Thr Val His Trp Ile Asp Arg Ala Thr Ile Thr Gly Tyr Arg Ile Arg 1360  
 His His Pro Glu His Phe Ser Gly Arg Pro Arg Glu Asp Arg Val Pro His Ser Arg Asn 1380  
 Ser Ile Thr Leu Thr Asn Leu Thr Pro Gly Thr Glu Tyr Val Val Ser Ile Val Ala Leu 1400  
 Asn Glu Arg Glu Glu Ser Pro Leu Leu Ile Gly Glu Glu Ser Thr Val Ser Asp Val Pro 1420  
 Arg Asp Leu Glu Val Val Ala Ala Thr Pro Thr Ser Leu Leu Ile Ser Trp Asp Ala Pro 1440  
 Ala Val Thr Val Arg Tyr Tyr Arg Ile Thr Tyr Gly Glu Thr Gly Gly Asn Ser Pro Val 1460  
 Glu Glu Phe Thr Val Pro Gly Ser Lys Ser Thr Ala Thr Ile Ser Gly Leu Lys Pro Gly 1480  
 Val Asp Tyr Thr Ile Thr Val Tyr Ala Val Thr Gly Arg Gly Asn Ser Pro Ala Ser Ser 1500  
 Lys Pro Ile Ser Ile Asn Tyr Arg Thr Glu Ile Asp Lys Pro Ser Glu Met Glu Val 1520  
 Asp Val Glu Asn Asn Ser Ile Ser Val Lys Trp Leu Pro Ser Ser Pro Val Thr Gly 1540  
 Tyr Arg Val Thr Thr Pro Lys Asn Glu Pro Gly Pro Thr Lys Thr Lys Thr Ala Gly 1560  
 Pro Asn Glu Thr Glu Met Thr Ile Glu Glu Leu Glu Pro Thr Val Glu Tyr Val Val Ser 1580  
 Val Thr Glu Asn Pro Ser Gly Glu Ser Glu Pro Leu Val Glu Thr Ala Val Thr Thr 1600  
 Ile Pro Ala Pro Thr Asn Leu Lys Phe Thr Glu Val Thr Pro Thr Ser Leu Ser Ala Glu 1620  
 Thr Thr Pro Pro Asn Val Glu Leu Thr Gly Tyr Arg Val Arg Val Thr Pro Lys Glu Lys 1640  
 Thr Gly Pro Met Lys Glu Ile Asn Leu Ala Pro Asp Ser Ser Val Val Val Ser Glu 1660  
 Leu Met Val Ala Thr Lys Tyr Glu Val Ser Val Thr Glu Leu Lys Asp Thr Leu Thr Ser 1680

FIGURE 5D

Arg Pro Ala Glu Gly Val Val Thr Thr <sup>1690</sup> Leu Glu Asn Val Ser Pro Pro Arg Arg Ala Arg  
 Val Thr Asp Ala Thr Glu Thr Thr Ile <sup>1710</sup> Thr Ile Ser Trp Arg Thr Lys Thr Glu Thr Ile  
 Thr Gly Phe Glu Val Asp Ala Val Pro Ala Asn Gly Glu Thr Pro Ile Glu Arg Thr Ile  
 Lys Pro Asp Val Arg Ser Tyr Thr Ile <sup>1730</sup> Thr Gly Leu Glu Pro Gly Thr Asp Tyr Lys Ile  
 Tyr Leu Tyr Thr Leu Asn Asp Asn Ala Arg <sup>1750</sup> Ser Ser Pro Val Val Ile Asp Ala Ser Thr  
 Ala Ile Asp Ala Pro Ser Asn Leu Arg Phe <sup>1770</sup> Leu Ala Thr Thr Pro Asn Ser Leu Leu Val  
 Ser Trp Glu Pro Pro Arg Ala Arg Ile <sup>1790</sup> Thr Gly Tyr Ile Ile Lys Tyr Glu Lys Pro Gly  
 Ser Pro Pro Arg Glu Val Val Pro Arg Pro <sup>1810</sup> Arg Pro Gly Val Thr Glu Ala Thr Ile Thr  
 Gly Leu Glu Pro Gly Thr Glu Tyr Thr Ile <sup>1830</sup> Tyr Val Ile Ala Leu Lys Asn Asn Glu Lys  
 Ser Glu Pro Leu Ile Gly Arg Lys Lys Thr <sup>1850</sup> Asp Glu Leu Pro Glu Leu Val Thr Leu Pro  
 His Pro Asn Leu His Gly Pro Glu Ile <sup>1870</sup> Asp Val Pro Ser Thr Val Glu Lys Thr Pro  
 Phe Val Thr His Pro Gly Tyr Asp Thr Gly <sup>1890</sup> Asn Gly Ile Glu Leu Pro Gly Thr Ser Gly  
 Glu Glu Pro Ser Val Gly Glu Glu Met Ile <sup>1910</sup> Phe Glu Glu His Gly Phe Arg Arg Thr Thr  
 Pro Pro Thr Thr Ala Thr Pro Ile Arg His <sup>1930</sup> Arg Pro Arg Pro Tyr Pro Pro Asn Val Ala  
 Leu Ser Glu Thr Thr Ile Ser Trp Ala Pro <sup>1950</sup> Phe Glu Asp Thr Ser Glu Tyr Ile Ile Ser  
 Cys His Pro Val Gly Thr Asp Glu Glu Pro Leu Glu Phe Arg Val Pro Gly Thr Ser Thr <sup>1970</sup>  
 2000 <sup>1980</sup>  
 Ser Ala Thr Leu Thr Gly Leu Thr Arg Gly Ala Thr Tyr Asn Ile Ile Val Glu Ala Leu  
 Lys Asp Glu Glu Arg His Lys Val Arg Glu <sup>2030</sup> Glu Val Val Thr Val Gly Asn Ser Val Asn  
 Glu Glu Leu Asn Glu Pro Thr Asp Asp Ser <sup>2050</sup> Cys Phe Asp Pro Tyr Thr Val Ser His Tyr  
 Ala Val Glu Asp Glu Trp Glu Arg Met Ser <sup>2070</sup> Glu Ser Gly Phe Lys Leu Leu Cys Glu Cys  
 Glu Ser Phe Glu Ser Glu His Phe Arg Cys Asp Ser Ser Arg Trp Cys His Asp Asn Glu <sup>2090</sup>  
 2100

FIGURE 5E

FNDEL1

Val Asn Tyr Lys Ile Gly Glu Lys Trp Asp Arg Gln Gly Glu Asn Gln Met Met Ser  
2110  
Val Asn Tyr Lys Ile Gly Glu Lys Trp Asp Arg Gln Gly Glu Asn Gln Met Met Ser 2120  
Cys Thr Cys Leu Gly Asn Gly Lys Glu Pro Lys Cys Asp Pro His Glu Ala Thr Cys  
2130  
Tyr Asp Asp Gly Lys Thr Tyr His Val Gly Glu Gln Trp Gln Lys Glu Tyr Leu Gly Ala  
2140  
Ile Cys Ser Cys Thr Cys Phe Gly Gly Gln Arg Gly Thr Arg Cys Asp Asn Cys Arg Arg  
2150  
Pro Ser Gly Glu Pro Ser Pro Glu Gly Thr Thr Ser Tyr Asn Gln Tyr Ser Gln  
2160  
Arg Thr His Gln Arg Thr Asn Thr Asn Val Asn Cys Pro Ile Glu Cys Phe Met Pro Leu  
2170  
2180  
Asp Val Gln Ala Asp Arg Glu Asp Ser Arg Glu 2190  
2200  
2210  
2220  
2230

FIGURE 5F

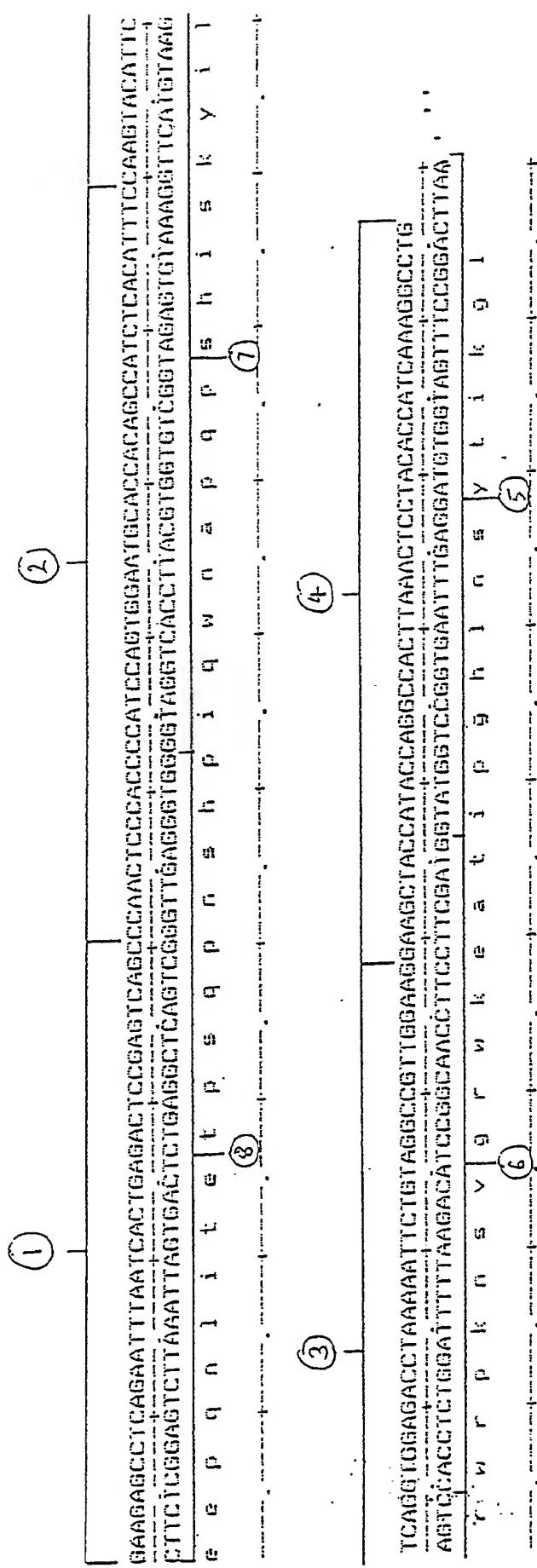


Figure 8. Linker 5 showing the eight constituent oligonucleotides.

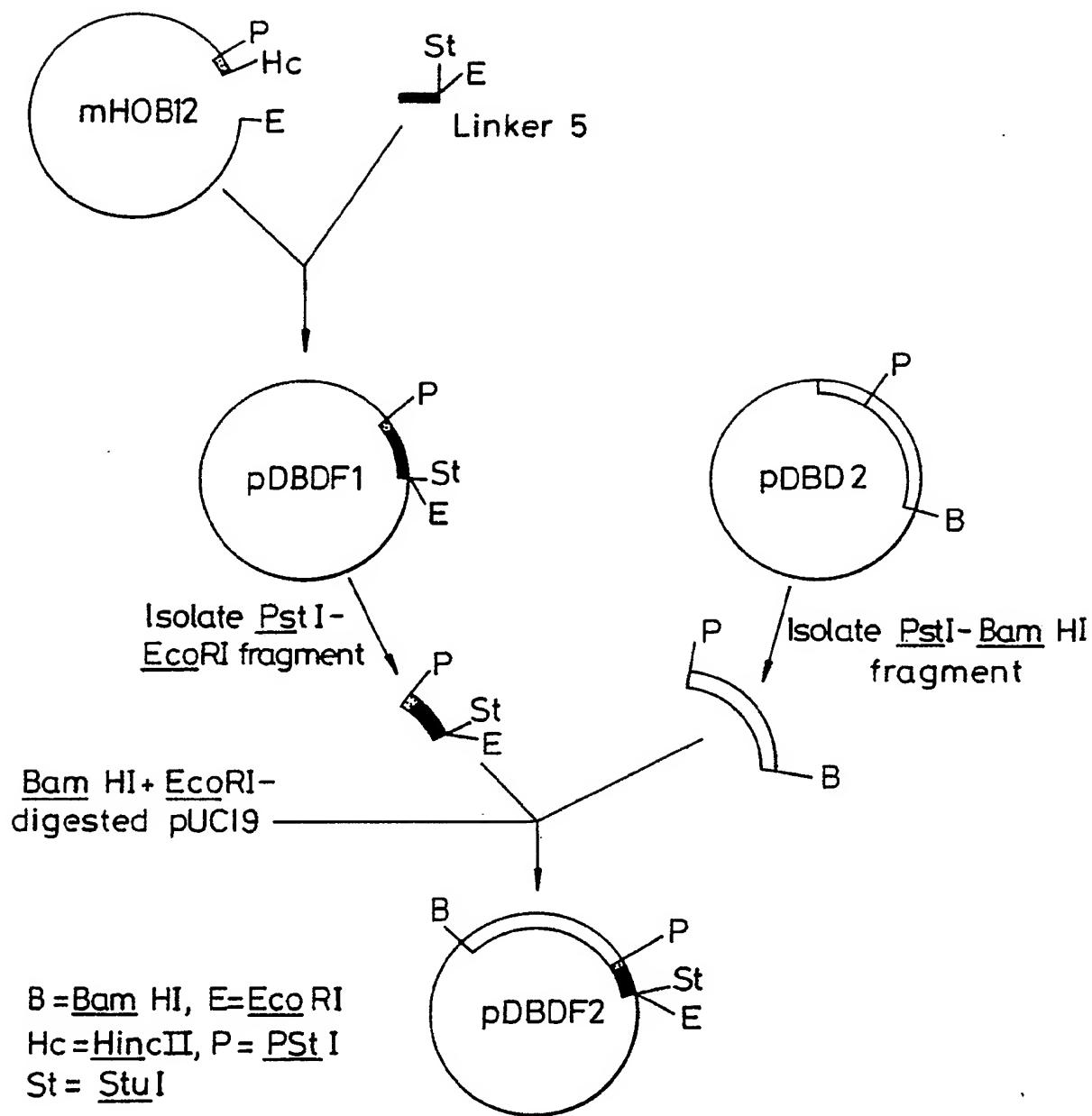


Fig. 7 Construction of pDBDF2

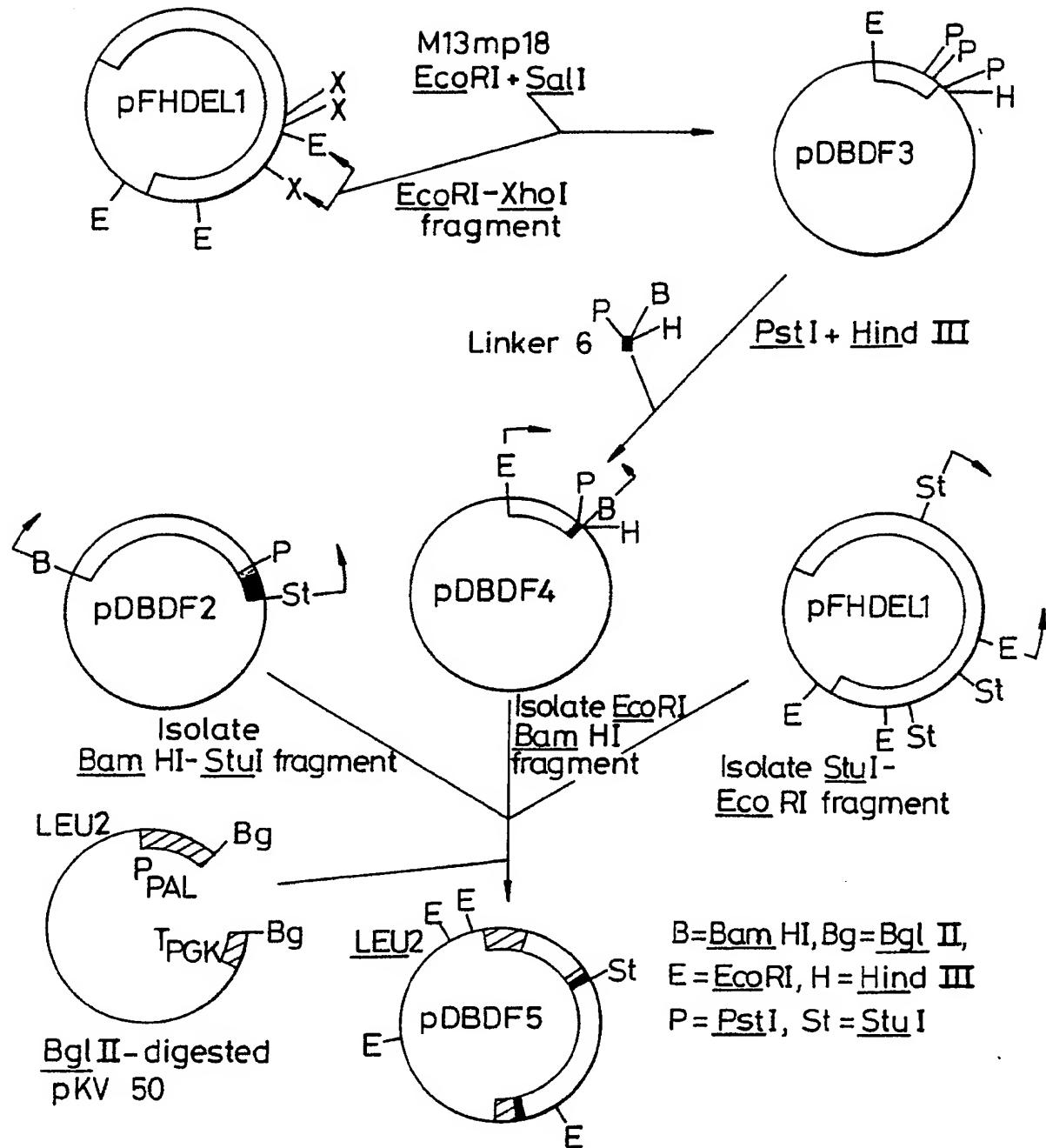


Fig. 8 Construction of pDBDF5

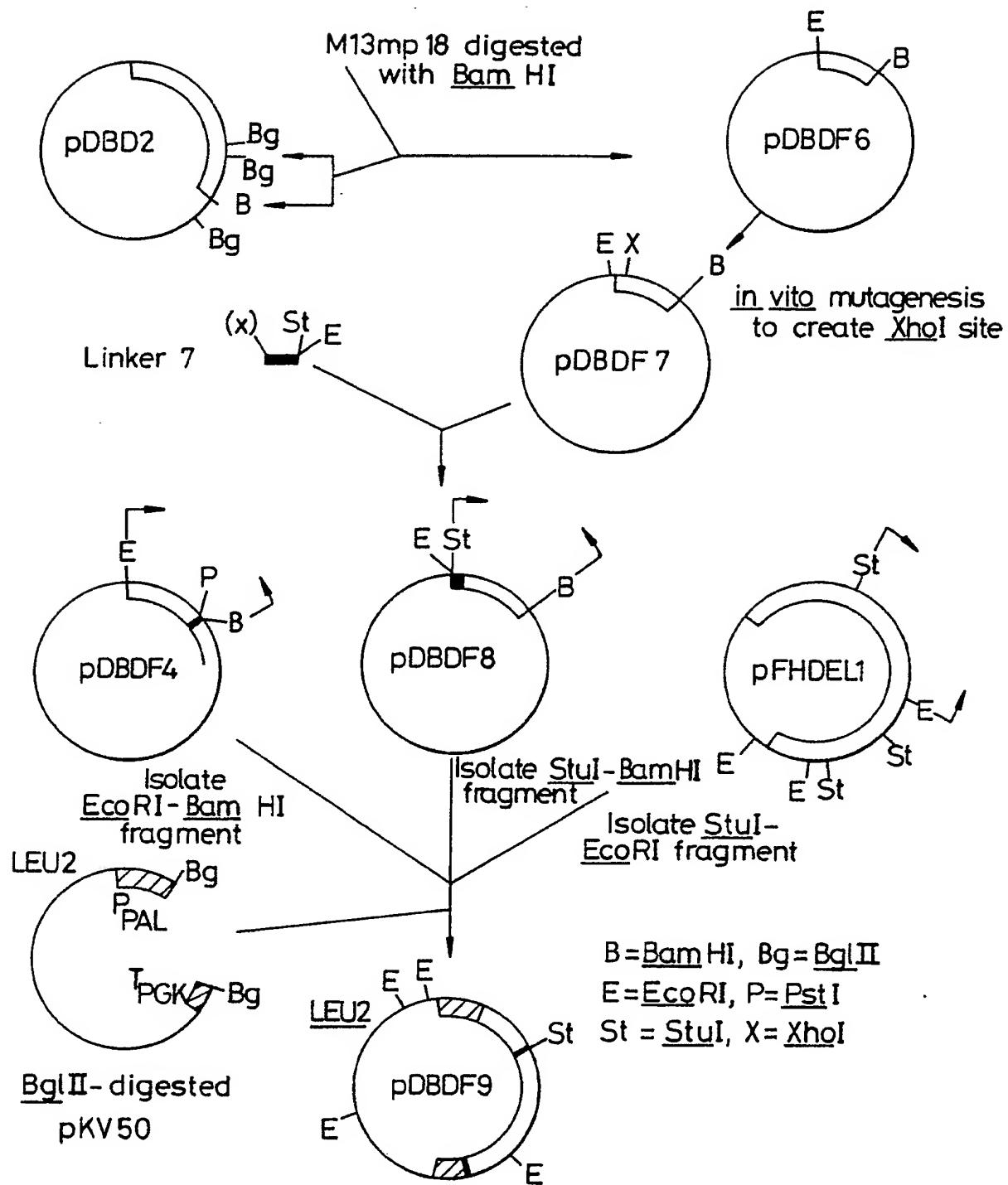


Fig. 9 Construction of pDBDF9

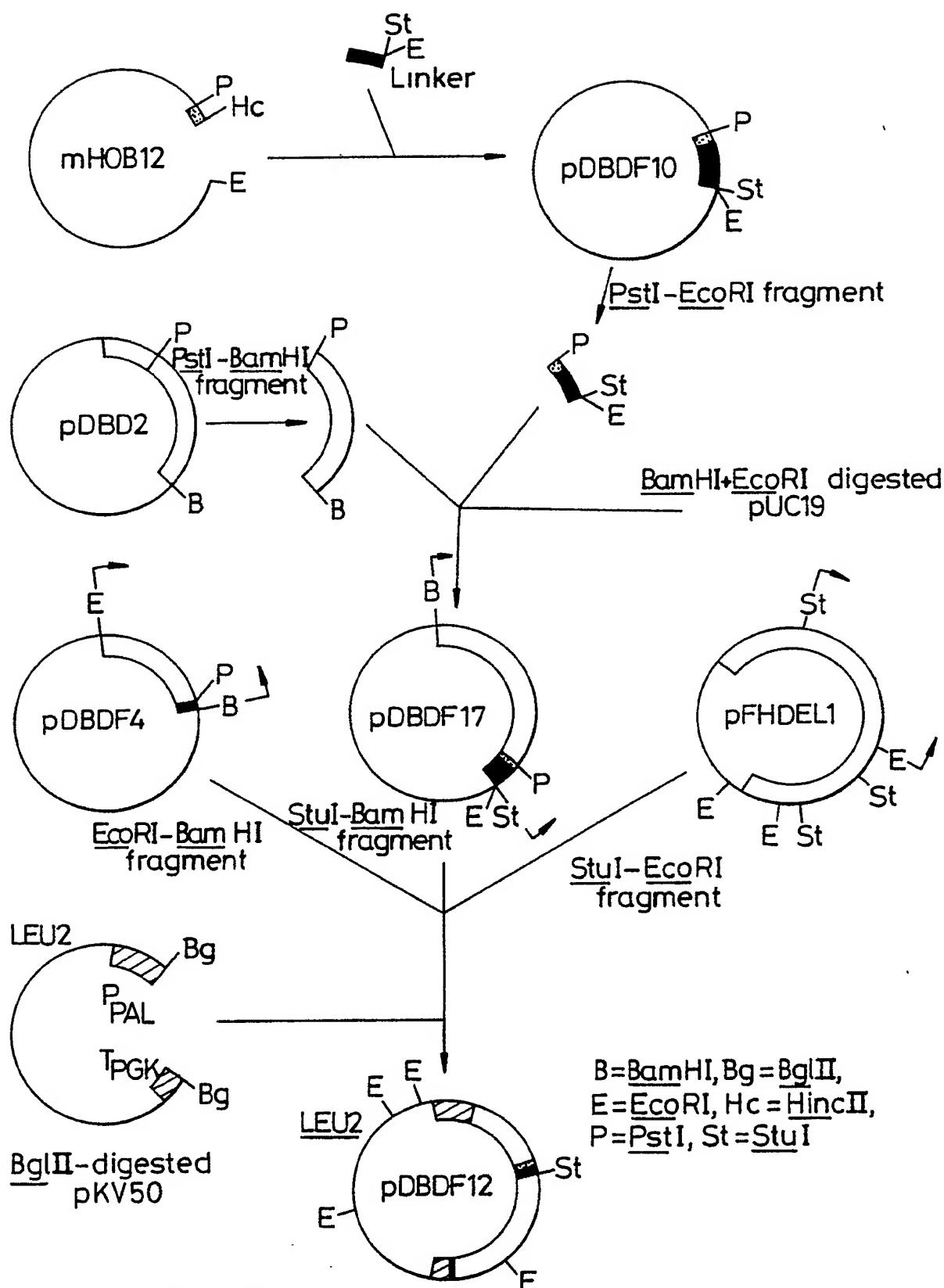
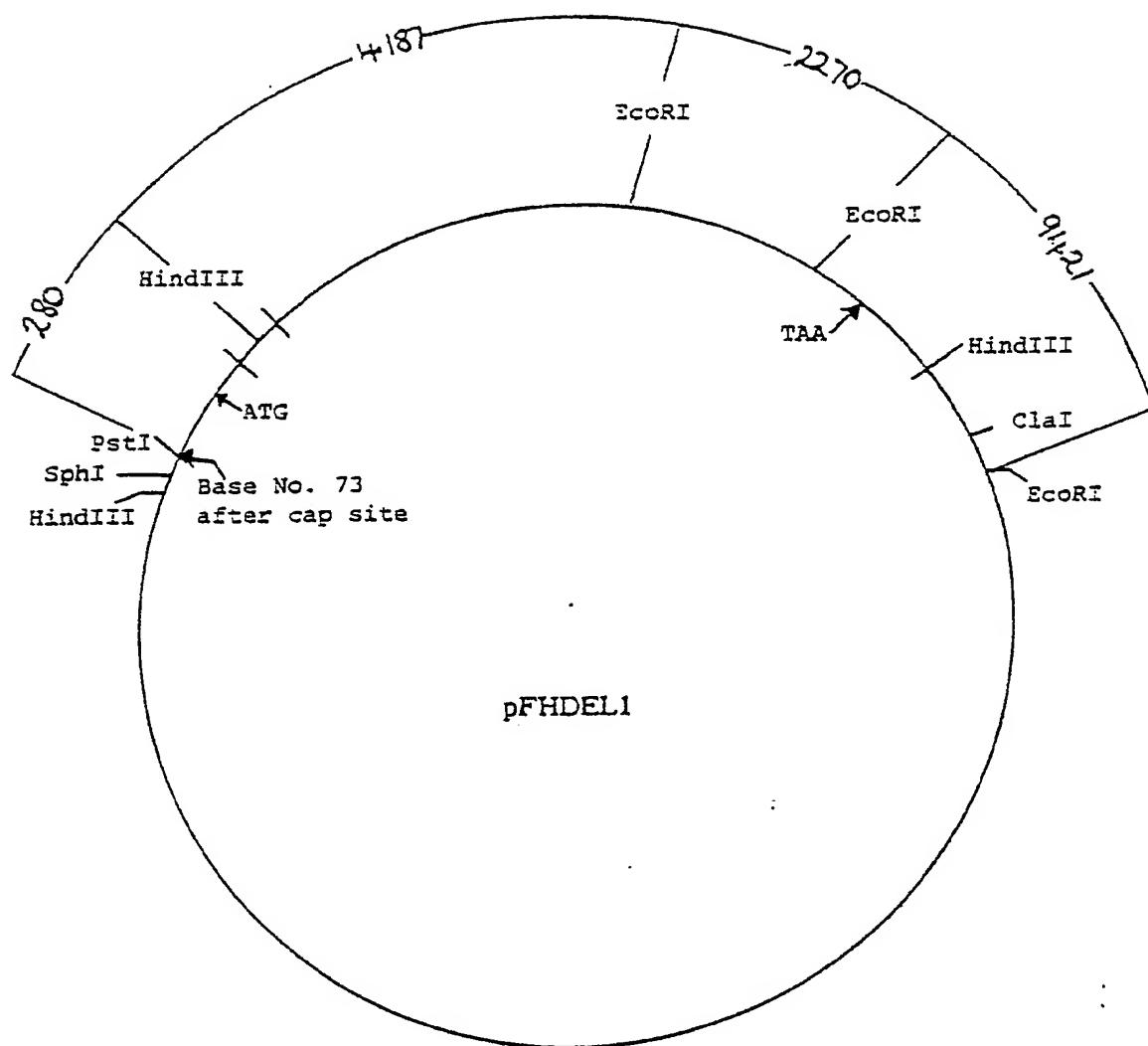


Fig. 10 Construction of pDBDF12

Figure 11

Name: pFHDEL1  
Vector: pUC18 Amp<sup>r</sup> 2860bp  
Insert: hFNCcDNA ~ 7630bp





European Patent  
Office

## EUROPEAN SEARCH REPORT

Application Number

EP 90 30 4575

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl. 5)
A	EP-A-0 308 381 (SKANDIGEN et al.) ---		C 12 N 15/62
T, D	EP-A-0 322 094 (DELTA BIOTECHNOLOGY LTD) -----		C 07 K 13/00 C 12 P 21/02
TECHNICAL FIELDS SEARCHED (Int. Cl. 5)			
C 12 N C 12 P C 07 K			
The present search report has been drawn up for all claims			
Place of search THE HAGUE	Date of completion of the search 10-07-1990	Examiner VAN PUTTEN A.J.	
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons ..... & : member of the same patent family, corresponding document	